Walmart Sales Forecasting and Analysis

Objective

This project aims to analyze historical sales data from Walmart stores, identify key trends, and build a time series forecasting model to predict future sales. The insights derived from this study can help in decision-making for inventory management, resource allocation, and overall business strategy.

Dataset Overview

The dataset consists of weekly sales data from multiple Walmart stores, including key economic indicators such as the **Unemployment Rate, Consumer Price Index (CPI), and Date-wise Sales Records**. The dataset also contains store-wise sales distributions, allowing a comparative analysis of top-performing and low-performing stores.

Exploratory Data Analysis (EDA)

- **Data Preprocessing**: Converted the date column into a standard format, checked for missing values, and handled outliers using statistical methods.
- **Statistical Summary**: Computed summary statistics to understand sales distribution, trends, and variations.
- **Correlation Analysis**: Examined the relationships between **Weekly Sales, CPI, and Unemployment Rate** to determine their impact on sales performance.
- **Seasonality Trends**: Analyzed sales patterns across different months and years to identify seasonal trends in sales.
- **Store Performance Analysis**: Identified the best and worst-performing stores based on total sales and visualized the comparison.

Time Series Forecasting using SARIMA Model

To predict future sales, we implemented a **Seasonal Autoregressive Integrated Moving Average (SARIMA)** model. The process involved:

1. **Checking Stationarity**: Used the **Augmented Dickey-Fuller (ADF) test** to determine whether the time series data was stationary. If the data was non-stationary, differencing was applied.
2. **Model Selection**: The **SARIMA (4,1,3)(1,1,0,52)** model was chosen based on the sales seasonality pattern (weekly sales data).
3. **Model Fitting**: Trained the model on historical sales data and evaluated the results.
4. **Forecasting**: Predicted sales for the next **12 weeks** and visualized the results with confidence intervals.

Key Findings & Insights

- **Sales seasonality**: Sales follow a **recurring pattern**, with higher sales during certain months, indicating seasonal demand.
- **Economic impact**: Unemployment and CPI have a moderate correlation with sales, suggesting external factors influence consumer spending.

- **Top-performing stores**: Identified stores generating the highest revenue, aiding in resource allocation strategies.
- **Sales forecasting**: The SARIMA model effectively predicts sales trends, providing insights for future business planning.

Conclusion

This project demonstrates how **data-driven decision-making** can enhance business operations in retail. By leveraging **machine learning, statistical modeling, and visualization techniques**, we extracted valuable insights that can help Walmart optimize inventory, marketing campaigns, and financial planning.